Dear Colleague:

At the recent Cold Spring Harbor retrovirus meeting it became apparent to some of us (especially those engaged in the rewriting of the CSH text and the preparation of various reviews) that the nomenclature for "transforming" inserts in retroviruses had become seriously confused. For example, seven different unrelated sequences, with completely different products were being called "*src*". It was also clear that if some agreement were not reached now, the situation would never be corrected. After much discussion, it was generally agreed that the best way to resolve the situation was to assign unique names to unrelated sequences according to the following rules:

1. As with other retroviral genes, the names should be 3 letters, lower case italics.

2. The names should be *trivial*; that is, no target cell specificity or functional significance is implied, and they are to be considered as names of coding sequences only.

3. They are to be derived in some mellifluous, yet mnemonic way from the name of the prototype virus or viruses, with the last letter suggestive of the major disease or other biological feature attributable to the sequence (for mnemonic reasons only).

4. Related sequences in different viruses are to be called by the same name, in a way that should (when completely resolved) point to the same cell sequence and the same or a closely related protein product, although it should not be necessary to have identified all of these to assign a name.

5. When necessary for clarity, the differences between inserts in related viruses can be indicated by using an "allele" designation, with a single letter indicating distinct isolates (e.g. mos^{G} vs mos^{M}), and additional letters or numbers to indicate strains (mos^{MI24}, src^{B77}, abl^{120}).

6. The related nonviral sequence found in the cell of origin will be designated with a lower case *c-* preceding the sequence name, e.g. *c-src*. The unadorned name will always indicate the viral sequence only. When necessary for emphasis, it can be prefixed with *v-*.

7. Protein products will be designated according to previous convention, e.g. $pp60^{src}$, $P150^{c\text{-}abl}$, $P110^{gag\text{-}abl}$.

8. Should the same virus be found to have two independently expressed inserts (i.e. coding for different proteins through distinct mRNAs) then two related but different names should be given.

9. Such names should be reserved for nonviral related sequences only. Such situations as SFFV and the 30S region of Ha and Ki MSV should not be so named.

10. Names along the same lines can also be given to nontransforming inserts if found in retroviruses or deliberately put there, but should be limited to genetically significant regions, i.e. those with protein (or functional RNA) products.

11. An exception to rule 4 can be made (although it need not) in the case where somewhat different yet related inserts are found in viruses of different species (after all, animal geneticists have also given different names to related loci found in different animals).

12. Strict genetic evidence is not required to assign a name, but it should be shown A) that the region is non-viral, and B) that it has either a protein (or functional RNA) product or a genetically identifiable function.

A list of the recommended names for known inserts is attached.

Thos who attended the discussion following the last session of the Cold Spring Harbor Meeting will recognize the rules and list as essentially as discussed there with only a few minor additions and changes.

We think all of us will recognize the need for such rules. Of late, it has become increasingly difficult to discuss, write, or teach about these viruses without additional unnecessary effort and confusion due only to poor or absent nomenclature. (How do you pronounce "sarc"? "src"? How can the mouse have more than 1 unique endogenous "sarc"? Isn't it rather awkward to keep writing "A-MuLV specific sequence" time and time again?) The system suggested is the only one we could come up with which satisfied what seemed to be the essential criteria: That the designations be consistent with previous convention; that they imply only a sequence with a product, not a disease or target; and that they readily distinguish virus from cell sequence and one virus from another.

It is our intention to publish these rules and the list of recommended names, signed by those workers agreeing to this nomenclature. If you agree

to use these names and are willing to have your name used, please let one of us know as soon as possible. We don't yet consider this formulation final (although we hope it is pretty close) and we welcome additional suggestions for changes in names and rules.

With best regards,

John M. Coffin
Department of Molecular Biology
and Microbiology
Tufts University School of Medicine
136 Harrison Avenue
Boston, Massachusetts 02111

Harold E. Varmus
Department of Microbiology
University of California at
San Francisco
Parnassus Avenue and 3rd Avenue
San Francisco, California 94143

JMC/evm